

**WEST**

Help

Logout

Interrupt

Main Menu

Search Form

Posting Counts

Show S Numbers

Edit S Numbers

Preferences

Cases

**Search Results -**

Terms	Documents
L9 and find\$ near4 pattern\$	39

US Patents Full-Text Database  
 US Pre-Grant Publication Full-Text Database  
 JPO Abstracts Database  
 EPO Abstracts Database  
 Derwent World Patents Index  
 IBM Technical Disclosure Bulletins

Database:

Search:

Refine Search

Recall Text

Clear

**Search History**DATE: Wednesday, March 05, 2003 [Printable Copy](#) [Create Case](#)

<u>Set Name</u> side by side	<u>Query</u>	<u>Hit Count</u>	<u>Set Name</u> result set
<i>DB=USPT,PGPB,JPAB,EPAB,DWPI,TDBD; PLUR=YES; OP=OR</i>			
<u>L10</u>	L9 and find\$ near4 pattern\$	39	<u>L10</u>
<u>L9</u>	query\$ near5 applications	3065	<u>L9</u>
<u>L8</u>	6119120.uref.	2	<u>L8</u>
<u>L7</u>	5564058.pn.	2	<u>L7</u>
<u>L6</u>	5774588.pn.	2	<u>L6</u>
<u>L5</u>	5706496.pn.	2	<u>L5</u>
<i>DB=USPT; PLUR=YES; OP=OR</i>			
<u>L4</u>	5459739.pn.	1	<u>L4</u>
<i>DB=USPT,PGPB,JPAB,EPAB,DWPI,TDBD; PLUR=YES; OP=OR</i>			
<u>L3</u>	6119120.pn.	2	<u>L3</u>
<u>L2</u>	4937763.pn.	2	<u>L2</u>
<u>L1</u>	5067099.pn.	2	<u>L1</u>

END OF SEARCH HISTORY

BEST AVAILABLE COPY

**WEST**

Generate Collection

Print

L10: Entry 27 of 39

File: USPT

Sep 18, 2001

US-PAT-NO: 6292797

DOCUMENT-IDENTIFIER: US 6292797 B1

TITLE: Method for determining actionable patterns in a database

DATE-ISSUED: September 18, 2001

## INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Tuzhilin; Alexander S.	New York	NY		
Adomavicius; Gediminas	Jersey City	NJ		

## ASSIGNEE-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY	TYPE CODE
New York University	New York	NY			02

APPL-NO: 09/ 130844 [PALM]

DATE FILED: August 6, 1998

## PARENT-CASE:

The present application claims the benefit, under 35 U.S.C. section 119(e), of U.S. Provisional Application No. 60/055,005, filed Aug. 7, 1997.

INT-CL: [07] G06 F 17/30

US-CL-ISSUED: 707/6; 707/3, 707/102, 707/203

US-CL-CURRENT: 707/6; 707/102, 707/203, 707/3

FIELD-OF-SEARCH: 707/6, 707/102, 707/10, 707/3, 707/203

PRIOR-ART-DISCLOSED:

## U.S. PATENT DOCUMENTS

Search Selected

Search ALL

	PAT-NO	ISSUE-DATE	PATENTEE-NAME	US-CL
<input type="checkbox"/>	<u>5325466</u>	June 1994	Kornacker	395/77
<input type="checkbox"/>	<u>5572604</u>	November 1996	Simard	382/224
<input type="checkbox"/>	<u>5581634</u>	December 1996	Heide	382/226
<input type="checkbox"/>	<u>5586240</u>	December 1996	Khan et al.	395/769
<input type="checkbox"/>	<u>5659743</u>	August 1997	Adams et al.	395/621
<input type="checkbox"/>	<u>5731986</u>	March 1998	Yang	364/491
<input type="checkbox"/>	<u>5764974</u>	June 1998	Walster et al.	707/6
<input type="checkbox"/>	<u>5774576</u>	June 1998	Cox et al.	382/160
<input type="checkbox"/>	<u>5794239</u>	August 1998	Walster et al.	707/6
<input type="checkbox"/>	<u>5809499</u>	September 1998	Wong et al.	707/6
<input type="checkbox"/>	<u>5832182</u>	November 1998	Zhang et al.	707/10
<input type="checkbox"/>	<u>5857169</u>	January 1999	Seide	704/256

## OTHER PUBLICATIONS

Hambaba, Intelligent Hybrid System for Data Mining, IEEE Catalog No. 96TH8177, p. 111, Mar. 1996.\*

Kamber et al. Generalization and Decision Tree Induction: Efficient Classification in Data Mining, IEEE, pp. 111-120, Apr. 1997.\*

Yongjlan, Data Mining, IEEE, pp. 18-20, 1997.\*

Tuzhilin et al., "A Belief-Driven Discovery Framework Based on Data Monitoring and Triggering," Center for Research on Inform Dec., 1996, pp. 1-23.

T. Imielinski et al., "DataMine: Application Programming Interface and Query Language for Database Mining", Systems for Mining Large Databases, KDD-96, pp. 256-261.

Han et al., "DMLQ: A Data Mining Query Language for Relational Databases", Database Systems Research Laboratory, pp. 27-33.

Agrawal et al., "Fast Discovery of Association Rules", pp. 307-328.

Klemettinen, "Finding Interesting Rules from Large Sets of Discovered Association Rules", University of Helsinki, pp. 1-7.

Silberchatz et al., "What Makes Patterns Interesting in Knowledge Discovery Systems", pp. 1-13.

Shen et al., "Metaqueries for Data Mining", pp. 375-397.

Matheus et al., "Selecting and Reporting What is Interesting: The Kefir Application to Healthcare Data", Advances in Knowledge Discovery and Data Mining, AAAI/MIT Press, 1995, pp. 401-419.

Agrawal et al., "Mining Association Rules between Sets of Items in Large Databases", IBM Almaden Research Center, pp. 207-216.

Silberschatz et al., "On Subjective Measure of Interestingness in Knowledge Discovery", pp. 275-281.

Piatetsky-Shapiro et al., "The Interestingness of Deviations", AAAI-94 Workshop on Knowledge Discovery in Databases, KDD-94, pp. 25-36.

ART-UNIT: 211

PRIMARY-EXAMINER: Black; Thomas

ASSISTANT-EXAMINER: Coby; Frantz

## ABSTRACT:

A user specifies a hierarchical action tree via user input device and user interface element. The action tree is arranged in a tree of file directories, with each node of the tree corresponding to a file directory (or path). The user then specifies classes of patterns assigned to each node (directory) of the tree using data mining queries or pattern templates. Once the system is so initialized, the pattern templates and data mining queries are executed, retrieving the patterns specified by the user from a database. The retrieved patterns assigned to a node of the tree are then stored in a file in the corresponding file directory. The user may now act on the discovered patterns and use the organized file structure. A pattern discovery optimization element periodically checks if the database has changed substantially, and if it has re-executes the data mining queries and pattern templates which update the contents of the file structure accordingly.

26 Claims, 5 Drawing figures

**WEST**

Generate Collection

Print

L10: Entry 38 of 39

File: USPT

Dec 10, 1996

US-PAT-NO: 5584024

DOCUMENT-IDENTIFIER: US 5584024 A

TITLE: Interactive database query system and method for prohibiting the selection of semantically incorrect query parameters

DATE-ISSUED: December 10, 1996

## INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Shwartz; Steven P.	Orange	CT		

## ASSIGNEE-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY	TYPE CODE
Software AG				DE	03

APPL-NO: 08/ 217099 [PALM]

DATE FILED: March 24, 1994

INT-CL: [06] G06 F 17/30, G06 F 17/27

US-CL-ISSUED: 395/604; 395/922, 395/757, 364/274.2, 364/275.4, 364/283.3, 364/DIG.1, 364/972.2, 364/974.6, 364/DIG.2

US-CL-CURRENT: 707/4; 704/7, 706/45, 706/922, 706/934

FIELD-OF-SEARCH: 395/600, 395/922, 364/419.01, 364/419.07, 364/974.6, 364/972.2, 364/274.2, 364/274.7, 364/275.1, 364/275.4, 364/283.3, 364/282.1

PRIOR-ART-DISCLOSED:

U.S. PATENT DOCUMENTS

Search Selected

Search ALL

PAT-NO	ISSUE-DATE	PATENTEE-NAME	US-CL
<input type="checkbox"/> 4506326	March 1985	Shaw et al.	364/300
<input type="checkbox"/> 4688195	August 1987	Thompson et al.	364/300
<input type="checkbox"/> 4689737	August 1987	Grant	364/200
<input type="checkbox"/> 4736296	May 1988	Katayama et al.	364/419
<input type="checkbox"/> 4811207	March 1989	Hikita et al.	364/200
<input type="checkbox"/> 4829423	May 1989	Tennant et al.	364/200
<input type="checkbox"/> 4839853	June 1989	Deerwester et al.	364/900
<input type="checkbox"/> 4914590	March 1990	Loatman et al.	364/419
<input type="checkbox"/> 4930071	May 1990	Tou et al.	364/300
<input type="checkbox"/> 4931935	May 1990	Ohira et al.	364/419
<input type="checkbox"/> 4943933	July 1990	Miyamoto et al.	364/513
<input type="checkbox"/> 4974191	November 1990	Amirghodsi et al.	364/900
<input type="checkbox"/> 4994967	February 1991	Asakawa	364/419
<input type="checkbox"/> 5099426	March 1992	Carlgren et al.	364/419
<input type="checkbox"/> 5175814	December 1992	Anick et al.	395/161
<input type="checkbox"/> 5197005	March 1993	Shwartz et al.	395/600
<input type="checkbox"/> 5204947	April 1993	Bernstein et al.	395/157
<input type="checkbox"/> 5237502	August 1993	White et al.	364/419
<input type="checkbox"/> 5255386	October 1993	Prager	395/600
<input type="checkbox"/> 5265014	November 1993	Haddock et al.	364/419
<input type="checkbox"/> 5265065	November 1993	Turtle	395/600
<input type="checkbox"/> 5349526	September 1994	Potts, Sr. et al.	364/419.1
<input type="checkbox"/> 5386556	January 1995	Hedim et al.	395/600

## FOREIGN PATENT DOCUMENTS

FOREIGN-PAT-NO	PUBN-DATE	COUNTRY	US-CL
0287310	October 1988	EP	
0387226	September 1990	EP	
63-219034	September 1988	JP	

## OTHER PUBLICATIONS

Wu, "A Knowledge-Based Database Assistant With A Menu Based Natural Language User-Interface" 10 Oct. 1993, IEICI: Trans. Inf. & Syst. V. E76-D N. 10 pp. 1276-1287.

Wu et al, "KDA: A Knowledge-Based Database Assistant With Query Guiding Facility" 5 Oct. 1992, pp. 443-453, IEEE Transactions On Knowledge & Data Eng. V. 4 N. 5.

Cha, "Kaleidoscope: A Cooperative Menu Guided Query Interface (SQL Version)," 1990, IEEE, Artificial Intelligence Applications.

Whitaker and Bonnell "Functional Modelling Of Intelligent Systems Using A Blackboard Model" 1992, pp. 1 to 12, The Journal Of Knowledge Eng. V 5, N. 1.

Winston, P., "Language Understanding," Artificial Intelligence, 9:291-334 (1984).

Rich, E., "Natural Lanugage Interfaces," Computer, pp. 39-47 (Sep. 1984).

Manferdelli, J. L., "Natural Language Interfaces: Benefits, Requirements, State of the Art and Applications," AI East, Oct. 1987.

Schank, R. C., et al, "Inside Computer Understanding: Five Programs Plus Miniatures," 14:354-372, LEA, Publishers, Hillsdale, NJ (1981).

Hendrix, G., "The Lifer Manual: A Guide to Building Practical Natural Language Interfaces" (Technical Note 138); SRI International, Feb. 1977.

Hendrix, G., "Human Engineering for Applied Natural Language Processing" (Technical Note 139); SRI International, SRI Project 740D32 CTC, Feb 1977.

Kao, M., et al, "Providing Quality Responses with Natural Language Interfaces: The Null Value Problem," IEEE 14:7, 959-984, Jul., 1988.

Chapter 6, "Queries," Building Access Z Applications, 1995.

Chapter 8, "Using Query by Example," Using Access 2 for Windows, Sp. Ed.  
Chapter 9, "Querying Your Data," Inside Paradox 5 for Windows, 1994.  
Cinque, L., et al, "An Expert Visual Query System," J. Vis. Lang. and Comp.,  
2:101-113 (1991).  
Meng, W., et al. "A Theory of Translation From Relational Queries to Hierarchical  
Queries," IEEE Transactions on Knowledge and Data Engineering, 7:2, 228-245, Apr.  
1995.  
Jakobson, G., et al, "CALIDA: A System for Integrated Retrieval from Multiple  
Heterogeneous Databases," Proceedings of the 3rd Int'l Conf. on Data and Knowledge  
Bases: Improving Usability and Responsiveness, Jun. 28-30, Jerusalem, Israel.

ART-UNIT: 237

PRIMARY-EXAMINER: Amsbury; Wayne

ASSISTANT-EXAMINER: Choules; Jack M.

ABSTRACT:

A database query system includes a query assistant that permits the user to enter only queries that are both syntactically and semantically valid (and that can be processed by an SQL generator to produce semantically valid SQL). Through the use of dialog boxes, a user enters a query in an intermediate English-like language which is easily understood by the user. A query expert system monitors the query as it is being built, and using information about the structure of the database, it prevents the user from building semantically incorrect queries by disallowing choices in the dialog boxes which would create incorrect queries. An SQL generator is also provided which uses a set of transformations and pattern substitutions to convert the intermediate language into a syntactically and semantically correct SQL query.

The intermediate language can represent complex SQL queries while at the same time being easy to understand. The intermediate language is also designed to be easily converted into SQL queries. In addition to the query assistant and the SQL generator, an administrative facility is provided which allows an administrator to add a conceptual layer to the underlying database making it easier for the user to query the database. This conceptual layer may contain alternate names for columns and tables, paths specifying standard and complex joins, definitions for virtual tables and columns, and limitations on user access.

27 Claims, 32 Drawing figures

**WEST**☐ **Generate Collection** **Print**

L3: Entry 1 of 2

File: USPT

Sep 12, 2000

US-PAT-NO: 6119120

DOCUMENT-IDENTIFIER: US 6119120 A

TITLE: Computer implemented methods for constructing a compressed data structure from a data string and for using the data structure to find data patterns in the data string

DATE-ISSUED: September 12, 2000

## INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Miller; John W.	Kirkland	WA		

## ASSIGNEE-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY	TYPE CODE
Microsoft Corporation	Redmond	WA			02

APPL-NO: 08/ 673427 [PALM]

DATE FILED: June 28, 1996

INT-CL: [07] G06 F 17/30

US-CL-ISSUED: 707/101; 707/6, 707/7, 707/3

US-CL-CURRENT: 707/101; 707/3, 707/6, 707/7

FIELD-OF-SEARCH: 382/229, 382/230, 382/231, 707/6, 707/3, 707/7, 707/101, 707/2

PRIOR-ART-DISCLOSED:

## U.S. PATENT DOCUMENTS

☐ **Search Selected****Search ALL**

PAT-NO	ISSUE-DATE	PATENTEE-NAME	US-CL
<input type="checkbox"/> 5459739	October 1995	Handley et al.	371/136

## OTHER PUBLICATIONS

"Dynamic Programming Alignment of Sequences Representing Cyclic Patterns", by Jens Gregor and Michael G. Thomason, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, No. 2, pp. 129-135, Feb. 1993.

"Searching Genetic Databases on Splash 2", by Dzung T. Hoang, Proceedings IEEE Workshop on FPGAs for Custom Computing Machines (Cat. No. 93TH0535-5), pp. 185-191, Apr. 5, 1993.

"Rapid-2, An Objecti-Oriented Association Memory Applicable to Genome Data Processing", by Denis Archambaud, Pascal Faudemay, and Alain Greiner Proceedings of the Twenty-Seventh Annual Hawaii International Conference on System Sciences, pp. 150-159, Jan. 1994.

"A Faster Algorithm Computing String Edit Distances", William J. Masek and Michael S. Paterson, Journal of Computer and System Sciences, 20, pp. 18-31, Aug. 6, 1979.

"Synthesis and Recognition of Sequences", by S.C. Chan and A.K.C. Wong, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 13, No. 12, pp. 1245-1255, Dec. 1991.

"Efficient Systolic String Matching", by G.M. Megson, Electronic Letters, vol. 26, No. 24, pp. 2040-2042, Nov. 1990.

ART-UNIT: 273

PRIMARY-EXAMINER: Au; Amelia

ASSISTANT-EXAMINER: Frederick, II; Gilberto

## ABSTRACT:

A method for constructing a data structure for a data string of characters includes producing a matrix of sorted rotations of the data string. This matrix defines an A array which is a sorted list of the characters in the data string, a B array which is a permutation of the data string, and a correspondence array C which contains correspondence entries linking the characters in the A array to the same characters in the B array. A reduced A' array is computed to identify each unique character in the A array and a reduced C' array is computed to contain every s.sup.th entry of the C array. The B array is segmented into blocks of size s. During a search, the A' and C' arrays are used to index the B array to reconstruct any desired row from the matrix of rotations. Through this representation, the matrix of rotations can thus be used as a conventional sorted list for pattern matching or information retrieval applications. A data structure containing only the A', B, and C' has very little memory overhead. The B array contains the same number of characters as the original data string, and can be compressed in a block wise manner to reduce its size. The A' array is a fixed size equal to the size of the alphabet used to construct the data string, and the C' array is variable size according to the relationship  $n/s$ , where n is the number of characters in the data string and s is the size of the blocks of the B array. Accordingly, the data structure enables a tradeoff between access speed and memory overhead, the product of which is constant with respect to block size s.

35 Claims, 8 Drawing figures



**WEST**

Help

Logout

Interrupt

Main Menu

Search Form

Posting Counts

Show S Numbers

Edit S Numbers

Preferences

Cases

**Search Results -**

Terms	Documents
L53 and list	5

BEST AVAILABLE COPY

Database:

US Patents Full-Text Database  
 US Pre-Grant Publication Full-Text Database  
 JPO Abstracts Database  
 EPO Abstracts Database  
 Derwent World Patents Index  
 IBM Technical Disclosure Bulletins

Search:

Refine Search

Recall Text

Clear

**Search History**
**DATE: Wednesday, March 05, 2003**   [Printable Copy](#)   [Create Case](#)
**Set Name**   **Query**  
 side by side

**Hit Count**   **Set Name**  
 result set

DB=USPT,PGPB,JPAB,EPAB,DWPI,TDBD; PLUR=YES; OP=OR

<u>L54</u>	L53 and list	5	<u>L54</u>
<u>L53</u>	L47 and attribute-value\$	6	<u>L53</u>
<u>L52</u>	L48 and attribute-value\$	1	<u>L52</u>
<u>L51</u>	L50 and attribute-value\$	1	<u>L51</u>
<u>L50</u>	L49 and list	102	<u>L50</u>
<u>L49</u>	L48 and query	134	<u>L49</u>
<u>L48</u>	L47 and (datamining or data with mining)	157	<u>L48</u>
<u>L47</u>	on-line near analytical near process\$ or olap	514	<u>L47</u>
<u>L46</u>	5564058.pn. and attribute-value	0	<u>L46</u>
<u>L45</u>	137 and 5781896.pn.	0	<u>L45</u>
<u>L44</u>	137 and 5774588.pn.	0	<u>L44</u>
<u>L43</u>	141 and 5774588.pn.	0	<u>L43</u>
<u>L42</u>	L41 and list	7	<u>L42</u>
<u>L41</u>	137 and pre-comput\$	7	<u>L41</u>
<u>L40</u>	L39 and query	2	<u>L40</u>

<u>L39</u>	l13 and attribute-value	3	<u>L39</u>
<u>L38</u>	L37 and attribute-value	1	<u>L38</u>
<u>L37</u>	L36 and least same value	113	<u>L37</u>
<u>L36</u>	l13 and great\$ same value	250	<u>L36</u>
<u>L35</u>	l13 and greatest-value	0	<u>L35</u>
<u>L34</u>	((((382/229)!.CCLS.) )	453	<u>L34</u>
<u>L33</u>	((((382/\$)!.CCLS.) )	34477	<u>L33</u>
<u>L32</u>	((((705/5)!.CCLS.) )	368	<u>L32</u>
<u>L31</u>	((((705/\$)!.CCLS.) )	18624	<u>L31</u>
<u>L30</u>	((((707/\$)!.CCLS.) )	14571	<u>L30</u>
<u>L29</u>	((707/1  707/2  707/3 )!.CCLS. )	5127	<u>L29</u>
<i>DB=USPT; PLUR=YES; OP=OR</i>			
<u>L28</u>	5819282.pn.	1	<u>L28</u>
<u>L27</u>	5878426.pn.	1	<u>L27</u>
<u>L26</u>	5966139.pn.	1	<u>L26</u>
<i>DB=USPT,PGPB,JPAB,EPAB,DWPI,TDBD; PLUR=YES; OP=OR</i>			
<u>L25</u>	L22 and attribute-value	1	<u>L25</u>
<u>L24</u>	L23 and list	10	<u>L24</u>
<u>L23</u>	L22 and attribute	14	<u>L23</u>
<u>L22</u>	L21 and query	16	<u>L22</u>
<u>L21</u>	L20 and (least same value or lesser same value)	23	<u>L21</u>
<u>L20</u>	L18 and find\$ same great\$ same value	42	<u>L20</u>
<u>L19</u>	L18 and find\$ near list	6	<u>L19</u>
<u>L18</u>	data adj mining or datamining	1959	<u>L18</u>
<u>L17</u>	L16 and (least same value or lesser same value)	67	<u>L17</u>
<u>L16</u>	L15 and find\$ same great\$ same value	105	<u>L16</u>
<u>L15</u>	((707/3 )!.CCLS. )	2540	<u>L15</u>
<u>L14</u>	L13 and values	636	<u>L14</u>
<u>L13</u>	find\$ near list	838	<u>L13</u>
<u>L12</u>	L4 and attribute-value\$ same list	1	<u>L12</u>
<u>L11</u>	L5 and attribute-value\$ same list	0	<u>L11</u>
<u>L10</u>	L6 and attribute-value\$ same list	0	<u>L10</u>
<u>L9</u>	L6 and attribute-value same list	0	<u>L9</u>
<u>L8</u>	L6 and attribute-value near list	0	<u>L8</u>
<u>L7</u>	L6 and attribute-value and list	7	<u>L7</u>
<u>L6</u>	L5 and comput\$	910	<u>L6</u>
<u>L5</u>	L4 and query	927	<u>L5</u>
<u>L4</u>	"data mining"	1941	<u>L4</u>
<u>L3</u>	"analytical mining"	0	<u>L3</u>
<u>L2</u>	"on-line analytical mining"	0	<u>L2</u>
<u>L1</u>	"on-line analytical mining" or olam	6	<u>L1</u>

END OF SEARCH HISTORY

**THIS PAGE BLANK (USPTO)**

**WEST**

Generate Collection

Print

L7: Entry 6 of 7

File: USPT

Mar 28, 2000

US-PAT-NO: 6044366

DOCUMENT-IDENTIFIER: US 6044366 A

TITLE: Use of the UNPIVOT relational operator in the efficient gathering of sufficient statistics for data mining

DATE-ISSUED: March 28, 2000

## INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Graffe; Goetz	Bellevue	WA		
Fayyad; Usama	Mercede Island	WA		
Chaudhuri; Surajit	Redmond	WA		

## ASSIGNEE-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY	TYPE CODE
Microsoft Corporation	Redmond	WA			02

APPL-NO: 09/ 029953 [PALM]

DATE FILED: March 16, 1998

INT-CL: [07] G06 F 17/30

US-CL-ISSUED: 707/2; 707/3, 707/4, 707/6, 707/102

US-CL-CURRENT: 707/2; 707/102, 707/3, 707/4, 707/6

FIELD-OF-SEARCH: 707/2, 707/3, 707/5, 707/6, 707/4, 707/102

PRIOR-ART-DISCLOSED:

## U.S. PATENT DOCUMENTS

Search Selected

Search ALL

	PAT-NO	ISSUE-DATE	PATENTEE-NAME	US-CL
<input type="checkbox"/>	<u>5201047</u>	April 1993	Maki et al.	707/4
<input type="checkbox"/>	<u>5511190</u>	April 1996	Sharma et al.	707/1
<input type="checkbox"/>	<u>5710915</u>	January 1998	McElhiney	707/3
<input type="checkbox"/>	<u>5713020</u>	January 1998	Reiter et al.	707/102
<input type="checkbox"/>	<u>5748905</u>	May 1998	Hauser et al.	709/249
<input type="checkbox"/>	<u>5819282</u>	October 1998	Hooper et al.	707/103
<input type="checkbox"/>	<u>5878426</u>	March 1999	Plasek et al.	707/102
<input type="checkbox"/>	<u>5966139</u>	October 1999	Anupam et al.	345/440

## OTHER PUBLICATIONS

Automating the Analysis and Cataloging of Sky Surveys, Advances in Knowledge discovery and Data Mining, Chapter 19, pp. 471-493, Fayyad, Usama M., Djorgovski, S. George and Weir, Nicholas, AAAI Press/The MIT Press, Menlo Park California; Cambridge, MA; London, England (1996).

SLIQ: A Fast Scalable Classifier for Data Mining, EDBT Conference, Mehta, Manish,

\*Agrawal, Rakesh and Rissanen, Jorma (1996).  
From Digitized Images to Online Catalogs. AI Magazine, pp. 51-66, Fayyad, Usama M., Djorgovski, S.G. and Weir, Nicholas, (Summer 1996).  
Data Mining and Knowledge Discovery: Making Sense Out of Data, Fayyad, Usama M.  
Submitted to IEEE Expert, for publication in the special issue on Data mining (Aug., 1996).  
Panning for Data Gold, New Scientist. No. 2031, Weekly E1-80, pp. 31-33 (May 25, 1996).

ART-UNIT: 277

PRIMARY-EXAMINER: Homere; Jean R.

ABSTRACT:

The invention concerns a method and apparatus for generating a tabulation of counts of occurrences of value combinations of a set of attributes over a relation consisting of a set of database records. The gathered counts (also referred to as sufficient statistics) of attribute occurrences or correlation counts is most preferably used in building a classification or density estimation model from the database records that can be used to predict some attribute values based on other attribute values. A new SQL operator designated the `UNPIVOT` operator operates by scanning the database records and for each record reorganizes that data to form an UNPIVOTED data record that include the combinations of attribute name, attribute value and the values for one or more selected class attributes. The UNPIVOTED table can be used to produce the desired sufficient statistics in one scan of the data using standard database engines. While materialization of UNPIVOTED table would cause a large added scan cost overhead, the UNPIVOT operator allows us to achieve the counts without the added cost by combining the UNPIVOT operator with other SQL `select` and `group by` operators the UNPIVOTED table can be counted without the need for materializing it. The result is a guaranteed one pass algorithm that does not incur the added scan cost factor. The savings in scan cost can extend to several orders of magnitude compared to other methodologies for getting the counts supported by current database engines. The sufficient statistics so gathered can be used to drive a variety of data mining algorithms.

21 Claims, 2 Drawing figures

**WEST****End of Result Set**

Generate Collection

Print

L25: Entry 1 of 1

File: USPT

Dec 10, 2002

US-PAT-NO: 6493637

DOCUMENT-IDENTIFIER: US 6493637 B1

TITLE: Coincidence detection method, products and apparatus

DATE-ISSUED: December 10, 2002

## INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Steeg; Evan W.	Kingston			CA

## ASSIGNEE-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY	TYPE CODE
Queen's University at Kingston	Kingston			CA	03

APPL-NO: 09/ 404714 [PALM]

DATE FILED: September 24, 1999

## PARENT-CASE:

This is a continuation of International Application PCT/CA98/00273, with an international filing date of Mar. 23, 1998, now abandoned, which claims benefits of the filing date of U.S. Provisional application 60/041472, filed Mar. 24, 1997.

INT-CL: [07] G06 F 17/30

US-CL-ISSUED: 702/19

US-CL-CURRENT: 702/19

FIELD-OF-SEARCH: 702/19

PRIOR-ART-DISCLOSED:

## FOREIGN PATENT DOCUMENTS

FOREIGN-PAT-NO	PUBN-DATE	COUNTRY	US-CL
2 283 840	May 1995	GB	

## OTHER PUBLICATIONS

Conner et al. Proc. Natl. Acad. Sci. USA, 80, 278-282, Jan. 1983.\*  
Steeg et al., "Introduction of Specific Point Mutations into RNA Polymerase II by Gene Targeting in Mouse Embryonic Stem Cells: Evidence for a DNA Mismatch Repair Mechanism," Proc. Natl. Acad. Sci. USA, vol. 87, No. 12, National Academy of Sciences USA, Jun. 15, 1990, pp. 4680-4684.  
R. Agrawal et al., "Fast Discovery of Association Rules," Chapter 12, Advances in Knowledge Discovery and Data Mining, pub. American Association for Artificial Intelligence, Menlo Park, California, .COPYRGT.1996, pp. 307-328.  
D. Altschuh et al., "Correlation of Co-ordinated Amino Acid Substitutions with Function in Viruses Related to Tobacco Mosaic Virus," J. Mol. Biol., vol. 193, 1987, pp. 693-707.  
R. Bahadur, "A Representation of the Joint Distribution of Responses to n Dichotomous Items," Chapter 9, Studies in Item Analysis, ed. H. Solomon, Stanford University Press, 1962, pp. 158-175.  
Carr et al., "Templates for Looking at Gene Expression Clustering," Statistical Computing & Statistical Graphics Newsletter, pp. 20-29 (Apr. 1997).  
A.F.W. Coulson et al., "Protein and nucleic acid sequence database searching: a

\*suitable case for parallel processing," The Computer Journal, vol. 30, No. 5, Oct. 1987, Cambridge, Great Britain, pp. 420-424.

U. Goebel et al., "Correlated Mutations and Residue Contacts in Proteins," Proteins, vol. 18, 1994, pp. 309-317.

R. Guigo et al., "Inferring Correlation between Database Queries: Analysis of Protein Sequence Patterns," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, No. 10, Oct. 1993, New York, New York, USA, pp. 1030-1041.

D. Heckerman, "Bayesian Networks for Knowledge Discovery," Advances in Knowledge Discovery and Data Mining, Chapter 11, pub. American Association for Artificial Intelligence, Menlo Park, California, .COPYRGHT.1996, pp. 273-306.

W. Hoeffding, "Probability Inequalities for Sums of Bounded Random Variables," Journal of the American Statistical Association, vol. 58, No. 301, Mar. 1963, pp. 13-30.

Klingler, T. et al., "Discovering Structural Correlations in .alpha.-Helices," Protein Science, vol. 3, 1994, pp. 1847-1857.

B. Korber et al., "Covariation of Mutations in the V3 Loop of Human Immunodeficiency Virus Type 1 Envelope Protein: an Information Theoretic Analysis," Proc. Natl. Acad. Sci. USA, vol. 90, Aug. 1993, pp. 7176-7180.

A. Krogh et al., "Hidden Markov Models in Computational Biology: Applications to Protein Modeling," J. Mol. Biol., vol. 235, 1994, pp. 1501-1531.

A.S. Lapedes et al., "Use of Adaptive Networks to Define Highly Predictable Protein Secondary-Structure Classes," Machine Learning, vol. 21, No. 1 / 2, Oct.-Nov. 1995, Boston, Massachusetts, USA, pp. 103-124.

C. de Marcken, "Unsupervised Language Acquisition," Ph.D. Thesis, M.I.T. (Sep. 1996) (title page, abstract, and pp. 82-93).

P. Michaud, "Clustering Techniques," Future Generation Computer Systems, vol. 13, Nov. 1997, NL, pp. 135-147.

R. Paturi et al., "The Light Bulb Problem," Information and Computation, vol. 117, No. 2, Mar. 1995, pp. 187-192.

Steeg, E. et al., "The Efficient Determination of Higher-Order Features in Protein Sequence Data" (Working Paper), Aug. 27, 1993, pp. 1-18. (FIGs. lost).

Steeg, E. et al., "Application of a Novel and Fast Information-Theoretic Method to the Discovery of Higher-Order Correlations in Protein Databases" in Proceedings of the 1998 Pacific Symposium on Biocomputing. Ed. Altman et al., World Scientific Publishing Co., New Jersey, pp. 573-584 (Jan. 4-9, 1998).

J. Williams et al., "A Process for Detecting Correlations between Dichotomous Variables," B. Kleinmütz, ed., Clinical Information Processing by Computer, Holt, Rinehart and Winston, New York, 1969, pp. 100-128.

ART-UNIT: 1631

PRIMARY-EXAMINER: Borin; Michael

#### ABSTRACT:

A method and system for detecting coincidences in a data set of objects, where each object has a number of attributes. Iteratively, equally-sized subsets of the data set are sampled, and coincidences (co-occurrences of a plurality of attribute values in one or more objects in the subset) are recorded. For each coincidence of interest, the expected coincidence count is determined and compared with the observed coincidence count; this comparison is used to determine a measure of correlation for the plurality of attributes for the coincidence. The resulting set of k-tuples of correlated attributes is reported, a k-tuple of correlated attributes being a plurality of attributes for which the measure of correlation is above a predetermined threshold. The method and system (implemented on an array of processing nodes) is suitable for protein structure analysis, e.g. in HIV research.

39 Claims, 22 Drawing figures

**WEST**

Generate Collection

Print

L54: Entry 4 of 5

File: USPT

Jun 29, 1999

US-PAT-NO: 5918232

DOCUMENT-IDENTIFIER: US 5918232 A

TITLE: Multidimensional domain modeling method and system

DATE-ISSUED: June 29, 1999

## INVENTOR-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY
Pouschine; Nicholas	Fremont	CA		
Stross; Kenner G.	Oakland	CA		
Brill; Michael L.	San Francisco	CA		

## ASSIGNEE-INFORMATION:

NAME	CITY	STATE	ZIP CODE	COUNTRY	TYPE CODE
Whitelight Systems, Inc.	Palo Alto	CA			02

APPL-NO: 08/ 978168 [PALM]

DATE FILED: November 26, 1997

INT-CL: [06] G06 F 17/30

US-CL-ISSUED: 707/103; 707/2, 707/3, 707/4

US-CL-CURRENT: 707/103R; 707/2, 707/3, 707/4

FIELD-OF-SEARCH: 707/103, 707/2, 707/3, 707/4

PRIOR-ART-DISCLOSED:

## U.S. PATENT DOCUMENTS

Search Selected

Search ALL

	PAT-NO	ISSUE-DATE	PATENTEE-NAME	US-CL
<input type="checkbox"/>	<u>5367619</u>	November 1994	Dipaolo et al.	395/149
<input type="checkbox"/>	<u>5495608</u>	February 1996	Antoshenkov	
<input type="checkbox"/>	<u>5560007</u>	September 1996	Thai	
<input type="checkbox"/>	<u>5592666</u>	January 1997	Perez	
<input type="checkbox"/>	<u>5664172</u>	September 1997	Antoshenkov	
<input type="checkbox"/>	<u>5666528</u>	September 1997	Thai	
<input type="checkbox"/>	<u>5742738</u>	April 1998	Koza et al.	395/13

ART-UNIT: 271

PRIMARY-EXAMINER: Amsbury; Wayne

ASSISTANT-EXAMINER: Lewis; Cheryl R.

ABSTRACT:



A system and method for computer modeling (10) and for creating hyperstructures (51) which are to be contained in a computer memory, which obtains measurements of physical objects and activities which are related to the entity to be modeled in the computer hyperstructure (51). The measurements are transformed into computer data which corresponds to the physical objects and activities external to the computer system (10). A plurality of independent dimensions (54) are created, where each dimension (54) includes at least one element (58). A plurality of cells (56) are created, each of which is associated with the intersection of two or more elements (58), each cell (56) being capable of storing at least one value. At least one rule domain (60) is associated with at least one cell (56), the rule domain (60) including at least one rule for assigning values to the associated cells (56). A domain modeling rule set (126) is prepared (300), which determines which of the rules will provide the value associated with each of the cells (56) wherein application of the domain modeling rule set (126) to the hyperstructure (51) causes a physical transformation of the data corresponding to said physical objects which are modeled in said hyperstructure (51).

Also disclosed is a method for querying computer hyperstructures (51), a Hyperstructure Query Language, and a "cell explorer", which allows direct viewing of the applied formulas that produce a specific value for a cell (56).

18 Claims, 17 Drawing figures